

EXHIBIT D

MORRISON FOERSTER

425 MARKET STREET
SAN FRANCISCO
CALIFORNIA 94105-2482

TELEPHONE: 415.268.7000
FACSIMILE: 415.268.7522

WWW.MOFO.COM

MORRISON & FOERSTER LLP
AUSTIN, BEIJING, BERLIN, BOSTON,
BRUSSELS, DENVER, HONG KONG,
LONDON, LOS ANGELES, MIAMI,
NEW YORK, PALO ALTO, SAN DIEGO,
SAN FRANCISCO, SHANGHAI, SINGAPORE,
TOKYO, WASHINGTON, D.C.

May 4, 2024

Writer's Direct Contact
+1 (213) 892-5656
ABennett@mofo.com

Via Email

Zach Savage
One Manhattan West
New York, NY 10001

Re: *The New York Times Company v. Microsoft Corporation, et al.*,
Case No.: 23-cv-11195-SHS

Counsel:

We write to respond to Plaintiff's letter regarding OpenAI's Responses and Objections to Plaintiff's First Set of Requests for Production. We are available to meet and confer on Monday, May 6, between 11:30 a.m.-1:30 p.m. ET, to discuss the below.

Global Issue: Rule 34(b)(2)(C)

To respond to Plaintiff's repeated reference to Rule 34(b)(2)(C) in its letter, OpenAI was clear in its responses about the scope of its production. Where it is producing documents, it says what it will produce. It is not searching for, collecting, or producing documents outside of the scope of what it has said it would produce. In any event, OpenAI's responses are not materially different than the Times' responses. If the Times believes OpenAI's responses are deficient, please explain why the Times' are not also deficient.

1. RFP Nos. 1, 3, 8, 14

OpenAI has agreed to make available for inspection "the pretraining data for models used for ChatGPT" because our understanding is that the issues in this litigation primarily concern the use of Plaintiff's works in the pre-training part of the training process. Compl. ¶¶ 83-97. To that point, the factual allegations in the complaint focus on the pre-training process. *Id.* There is not a single fact supporting the allegation that post-training or fine-tuning is relevant to the issues in this case (although we acknowledge Plaintiff makes a few conclusory, unsupported assertions to that effect). Plaintiff has not articulated why post-training is relevant. To the extent Plaintiff is willing to articulate relevance, we are willing to meet and confer about it.

|||ORRISON FOERSTER

Consequently, we continue to believe that permitting inspection of the relevant pretraining data for models used for ChatGPT is sufficient to satisfy RFP Nos. 8 and 14. We are willing to meet and confer on the responses as to RFP Nos. 1 and 3 to identify an appropriate scope, although we believe search terms and custodians will likely be the appropriate method for identifying relevant documents.

2. RFP Nos. 5, 12, 18

Generative AI Products and Services

Plaintiff's definition of "Generative AI Products and Services" includes non-OpenAI, third-party products. RFPs including this term thus seek discovery into OpenAI's documents about third-party products. Copilot is not an OpenAI product. Nor are Bing Chat or "other platforms otherwise providing access to ChatGPT" that were built by third parties with ChatGPT API. OpenAI built a discrete set of large language models and used those models to build ChatGPT. OpenAI will provide Plaintiff with responsive information in its possession, custody, or control regarding its own products and services—the OpenAI-created models and ChatGPT. We do not understand why OpenAI's documents regarding products built by third parties are relevant.

As to Browse with Bing, we do not believe there is actually a dispute between the parties. Because Browse (formerly known as "Browse with Bing") is a feature within ChatGPT, our definition of "Generative AI Products and Services" to mean "ChatGPT" encompasses the Browse feature.

AI Models

We limited our definition of "AI Models" to mean only models used for ChatGPT for the same reasons as described above.

3. RFP No. 5

As we objected in our response, we do not know what you mean by "the use of intellectual property in AI models." We responded to the best of our understanding by agreeing to provide documents sufficient to explain the "Copyright Shield," which was explicitly identified in your response. We are willing to meet and confer to better understand what documents Plaintiff is looking for beyond what we have agreed to produce.

Additionally, in your letter you provide a link Footnote 1. The link requires a subscription, and thus we are unable to assess its content. OpenAI will therefore not consider it nor the arguments made by the Times in reliance on it. We are willing to meet and confer to discuss.

HARRISON FOERSTER**4. RFP Nos. 12, 13**

First, for the reasons described in Part 2.a above (regarding the definition of “Generative AI Products and Services”), there is no basis for Plaintiff to request documents concerning RAG’s use in products developed by third parties. We do not understand what such products and services RFP No. 12 is directed to other than RAG’s use in ChatGPT.

Second, Plaintiff provides no meaningful basis for its claim that OpenAI’s willingness to produce documents “sufficient to show” how OpenAI uses RAG in connection with ChatGPT (RFP No. 12) or its understanding of RAG (RFP No. 13) is “inappropriate.” OpenAI has agreed to produce a set of responsive documents, despite the two plainly overbroad RFPs that ask generally about a common technological technique employed across the AI industry. If Plaintiff has authority supporting its contention, please share it so we can evaluate it.

5. RFP No. 2

We understand that the FTC’s focus is on privacy and data security practices, and practices that allegedly resulted in “reputational harm” to consumers, not copyright infringement. Thus, the documents collected to be reviewed and submitted to the FTC are not relevant to this copyright infringement action. Plaintiff has not made any attempt to demonstrate why those interrogatories are relevant to its claims. *Citizens Union of N.Y.C. v. Att’y Gen. of N.Y.* 269 F. Supp. 3d 124, 139 (S.D.N.Y. 2017) (“The party seeking discovery bears the initial burden of proving the discovery is relevant.”). To the extent that there may be any overlapping documents relevant to this matter, those will be identified in a less burdensome, more efficient manner via the custodial and non-custodial collections in this matter. Requesting OpenAI to search and produce the documents collected in connection with the FTC investigation (which include not only non-custodial collections, but also custodial collections) would be inappropriate, particularly given the limited—if any—relevance, and the more appropriate discovery methods available to Plaintiff. *See In re CI Invs., Inc.*, No. 23 Misc. 434 (GHW) (GS), 2023 WL 8643965, at *6 (S.D.N.Y. Dec. 14, 2023) (citing cases); *N.J. Carpenters Health Fund v. DLJ Mortg. Cap., Inc.*, No. 08 Civ. 5653 (PAC), 2012 WL 13135408, at *1 (S.D.N.Y. Mar. 2, 2012) (denying discovery request characterized by the court as “you gave some documents to the government concerning another investigation, so give them to me”); *Humana Inc. v. Teva Pharms. USA, Inc.*, No. 6:21-CV-72-CEM-DCI, 2021 WL 8651837, at *1 (M.D. Fla. Oct. 19, 2021) (“Generally, cloned discovery has not been compelled, even absent a showing of undue burden, unless the requesting party shows that all the cloned discovery is relevant.”).

6. RFP No. 4

As we explained in our response, we believe the terms “claim” and “complaint” are vague. Nonetheless, in an effort to be responsive and as a reasonable interpretation of these

III HARRISON FOERSTER

words,¹ we proposed “formal, written complaint.” Although in its letter Plaintiff purports to seek more than such “complaints,” your letter still uses the term “any claim of copyright infringement.” To the extent that Plaintiff seeks information beyond a formal, written complaint, OpenAI requests that Plaintiff clarify what it means by “claims of copyright and trademark infringement, including complaints regarding infringement.”

7. RFP No. 6*Agreements*

OpenAI agreed to produce agreements with third parties relating to training data that was used to train the models for ChatGPT. We believed that such agreements would be relevant and responsive to the request for documents “concerning alternatives to using copyrighted content to train,” in part because Plaintiff referenced agreements. Notably, Plaintiff has objected to producing the type of materials it seeks here—namely, correspondence relating to such agreements. Please explain why Plaintiff now believes such communications are relevant.

Public Domain

As we stated in our objections, we do not understand what Plaintiff means by “alternatives to using copyrighted content.” We believed we understood the request to refer to agreements with third parties relating to training data, and agreed to produce such agreements. But we do not understand what you mean about using “material in the public domain” as an alternative to using copyrighted content. We are willing to meet and confer about the scope of the response.

8. RFP No. 7

To the extent that documents exist reflecting OpenAI’s understanding of ChatGPT’s effect, if any, on journalism, OpenAI has agreed to produce them. This includes OpenAI’s understanding regarding the use of ChatGPT as a source of, substitute for, or product in competition with journalism (as understood by OpenAI). Thus, your letter raises a dispute that appears to not exist.

In addition, the Times appears to seek documents that reflect *third parties’* understanding of the effect, if any, of ChatGPT on journalism. OpenAI is willing to meet and confer to understand the basis for the Times’ demand, given this information appears equally available to Plaintiff.

¹ In general, a claim is a “legal assertion.” See Black’s Law Dictionary (definition of “claim”).

III HARRISON FOERSTER

9. RFP No. 11

Plaintiff's RFP is based on a false premise. None of the OpenAI entities has transitioned to a for-profit company. Rather than object, we agreed to provide the documents necessary to show Plaintiff that no OpenAI entity has undergone the alleged "changes to the for-profit status." Because OpenAI has not "transitioned to a for-profit company," there are no documents that concern any "transition." OpenAI will not be producing additional documents in response to this request.

10. RFP 15

Plaintiff's request is overbroad on its face. It requests any "commercial . . . collaborations between and among" Microsoft and OpenAI. Because this seeks documents irrelevant to the claims and defenses of this litigation, we agreed to meet and confer regarding an appropriately narrow scope. To be clear, it is Plaintiff's burden to demonstrate relevance. OpenAI will not discuss the types of documents available until Plaintiff makes that facial showing. We look forward to hearing from Plaintiff for our consideration the relevance of the broad category of documents it seeks from this request, including a specific identification of how it makes element(s) of a claim more or less likely.

Sincerely,

/s/ Allyson R. Bennett

Allyson R. Bennett